

# A POMDP APPROACH TO UNDERWATER ROBOT PATH PLANNING FOR MULTI-VIEW MULTI-TARGET CLASSIFICATION

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Quanxing Lu

December 2018

© 2018 Quanxing Lu  
ALL RIGHTS RESERVED

# A POMDP APPROACH TO UNDERWATER ROBOT PATH PLANNING FOR MULTI-VIEW MULTI-TARGET CLASSIFICATION

Quanxing Lu

Cornell University 2018

This thesis presents an approach of classifying multiple targets of interest in minimum time with satisfactory confidence by an imaging sensor on an underwater robot. The overall goal is achieved by sequentially solving a single target classification problem and a global target ordering problem. First, a multi-view single-target classification algorithm is developed based on the POMDP framework, which incorporates a deep convolutional neural network and a support vector machine as the observation model. The classification algorithm allows the underwater robot to adaptively select its next configuration state near the target of interest in order to maximize the increase of classification confidence. Next, a traveling salesman algorithm is used to generate the global target visiting order. Simulation results of an unmanned underwater vehicle equipped with a side-scan sonar validate the effectiveness of the proposed algorithm and demonstrates the ability to find significantly shorter path for multi-view based multi-target classification.

## **BIOGRAPHICAL SKETCH**

Quanxing Lu is an M.S. student in the Mechanical and Aerospace Engineering Department at Cornell University. He received the B.S. degree in Mechanical Engineering from the University of California, San Diego. He has been working with Dr. Silvia Ferrari in the Laboratory of Intelligent Systems and Controls (LISC) since October 2016. His research interests include neural networks, machine learning, and optimal control.

## ACKNOWLEDGEMENTS

First, I would like to express my deeply gratitude to my advisor, Dr. Silvia Ferrari, for her helpful advice and enormous support during my graduate study at Cornell University. It was my great honor to work on this research project under her supervision. I would like to thank Dr. Douglas MacMartin for being my graduate committee member and offering his valuable suggestions on my research. I also truly appreciate Dr. Pingping Zhu, Dr. Bo Fu, and Dr. Chang Liu at LISC, who have also provided tremendous help in my research and suggestions in my preparation for my thesis defense. Last but not the least, I am grateful to all the LISC members and students at Cornell University, who have given me constructive feedback and great support during the work. This work was funded by ONR grant N00014-15-1-2595.

## CONTENTS

Biographical Sketch . . . . .	iii
Acknowledgements . . . . .	iv
Contents . . . . .	v
List of Tables . . . . .	vi
List of Figures . . . . .	vii
<b>1 Introduction</b>	<b>1</b>
<b>2 Problem Formulation</b>	<b>4</b>
<b>3 General POMDP Formulation</b>	<b>6</b>
<b>4 Image Template Issue</b>	<b>9</b>
<b>5 POMDP-based underwater robot path planning</b>	<b>13</b>
5.1 Observation Model . . . . .	13
5.2 C-target formation and Waypoints Selection . . . . .	14
5.3 POMDP Supplement . . . . .	15
5.4 Target ordering using TSP algorithm . . . . .	17
<b>6 Numerical Results</b>	<b>18</b>
6.1 CNN and SVM performace . . . . .	18
6.2 Observation Model Validation . . . . .	19
6.3 Simulation Results . . . . .	21
<b>7 Conclusion and Future Work</b>	<b>27</b>
<b>Bibliography</b>	<b>28</b>

## LIST OF TABLES

6.1	Classification performance of the CNN and SVM confusion matrix	18
6.2	Conditionally probability table of observation model . . . . .	19
6.3	The conditional probability table of test data when the ground truth $y_i$ is 1 (target) . . . . .	19
6.4	The conditional probability table of test data when the ground truth $y_i$ is 0 (non-target) . . . . .	19
6.5	The chi square value when the ground truth $y_i$ is 1 (target) . . . .	20
6.6	The chi square value when the ground truth $y_i$ is 0 (non-target) .	20
6.7	Performance comparison of the proposed algorithm and two MAC algorithms . . . . .	25

## LIST OF FIGURES

2.1	Side-scan and forward-looking sonar investigating underwater environment and corresponding sonar image . . . . .	4
4.1	Cylinder target from several aspects and ranges with the corresponding matching templates (Myers, 2012) . . . . .	10
4.2	Wedge target from several aspects and ranges with the corresponding matching templates (Myers, 2012) . . . . .	10
4.3	Segmented sonar images of rectangle shape target from several different regions . . . . .	11
4.4	Segmented sonar images of circle shape target from several different regions . . . . .	12
5.1	Work flow chart of PODMP-based underwater robot path planning	13
5.2	The discretization of the C-target based on the sensor range and aspect angle, which is defined as a region, and vehicle waypoints $K_1, K_2, K_3$ . . . . .	14
6.1	Robot trajectory generated by proposed algorithm . . . . .	22
6.2	An example of vehicle paths for a POMDP policy in the data set. The regions visited are determined by the policy, and the waypoints are set to allow the vehicle to follow a straight line that will get a good quality sonar image. . . . .	23
6.3	An example of vehicle paths for a POMDP policy in the data set.	23
6.4	An example of vehicle paths for a POMDP policy in the data set.	24
6.5	Robot trajectory generated by MAC algorithm . . . . .	25
6.6	Robot trajectory generated by MAC algorithm with clustering . .	26



## CHAPTER 1

### INTRODUCTION

The problem of autonomous target classification with imaging sensors has the potential to play a critical role in a wide range of applications. An autonomous robot equipped with onboard imaging sensors are now routinely employed for underwater, ground, and aerial target classification task. In many instances, although prior measurements may be available, accurate target classification requires images obtained from multiple view points in order to overcome environmental uncertainty and clutter. One approach that has been proposed in the literature [1] introduces frameworks that are based on either the Dempster-Shafer (DS) concept of fusion from a single-view classifier or direct multi-instance classification. In [2], the measure of similarity was used in order to classify the shape of a target by comparing set of views of a test target to a set of views from a training target whose the shape is known. The shape of maximum similarity was then used for classification obtaining significant performance improvements as the number of views of the test target increased. In [3], a multilayer perceptron (MLP) was used to enhance a hidden Markov model (HMM), and the enhanced model was used to classify underwater targets from multiple aspects.

The approach presented in this thesis is inspired by the partially observable Markov decision process (POMDP) model presented in [4] and [5], for the classification of a single underwater target using sonar images obtained from multiple view points. The method in [5] allows the vehicle to move around the target adaptively to acquire additional views that are expected to mostly reduce classification uncertainty. The POMDP approach in [5] was demonstrated via

synthetic aperture sonar (SAS) high-resolution images and, thus, an observation model was based on the similarity between the sonar image and a set of pre-generated image templates. However, the use of pre-generated templates limits the generalization capabilities of the observation model and, thus, the applicability of the view planning algorithm. Also, for different applications, the image templates need to be re-designed by the experts in order to capture the characteristics of the chosen sensor application, domain, and targets. Moreover, template-matching may become inefficient when the environmental conditions are complex. Meanwhile, low image resolution decreases the classification performance because it is typically challenging to match measurements to templates. Finally, the approach in [5] does not extend to the case of multiple targets, since the view point is selected incrementally, without the ability to optimize the robot path or avoid obstacles.

The approach presented in this thesis overcomes the aforementioned limitations by learning a POMDP and observation model from a convolutional neural network (CNN) and support vector machine (SVM) classifier [6] [7] trained with a database of underwater images obtained via side-scan sonar. Adopting the approach previously presented by the authors in [8], a pre-trained CNN AlexNet [9] is applied to extract convolutional features from segmented target sonar images. CNN is one of the deep learning architect that has a satisfying performance at visual image feature extraction [10] [11] [12] [13] [14]. The convolutional features are then used as inputs to an SVM [15] [16], thus making it possible to utilize the power of CNN with a relatively small database of target images. Subsequently, the approach presented by the authors in [17] is adopted for fusing the observations from multiple target images via CNN and SVM, such that the *posterior* distribution for target features and classification

are updated by probability inference based on image obtain via multiple view points for each target. The approach in [17] is utilized to provide an estimate of classification confidence level that can be used to adaptively select the next robot configuration with which to obtain the next best image.

This thesis presents a multi-aspect classification approach that obtains a POMDP model of a robot-based sensor from the image features extracted from CNN and SVM. The approach presented in this thesis adopts a traveling salesman problem (TSP) algorithm [18]. The traveling salesman problem (TSP) can be described as a traveling salesman starts at his home city and need to visit all cities around exactly once a time and return to the home city with a minimum time [19] [20]. It represents a large class of problems called combinatorial optimization problems and has a wide range of applications such as path planning [21] [22]. The problem is relatively hard to solve and in this thesis, the cross-entropy (CE) algorithm introduced in [23] and [24], is used to find an approximate solution. The combination of the POMDP mode and the TSP allow the proposed algorithms to plan the optimal robot path while adaptively selecting the next robot configuration so as to maximize the target classification confidence.

In the next section, the multi-view, multi-target classification problem is formulated. Then, the general POMDP framework is reviewed in Chapter 3. The proposed multi-view multi-target classification algorithm is described in Chapter 4. Finally, validation and performance evaluation is shown in Chapter 5.

## CHAPTER 2

### PROBLEM FORMULATION

Consider the problem of classifying  $N$  fixed targets in an underwater ROI,  $\mathcal{W} \in \mathbb{R}^2$ . An unmanned underwater vehicle (UUV) is deployed at the chosen starting configuration with a side scan sonar to classify targets as shown in Figure 2.1.

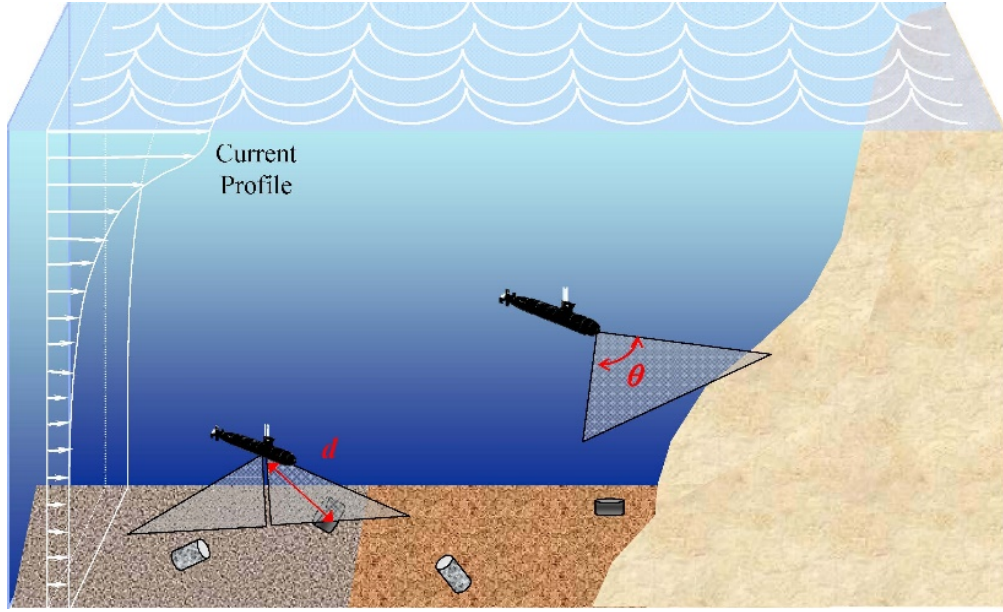


Figure 2.1: Side-scan and forward-looking sonar investigating underwater environment and corresponding sonar image

The onboard side-scan sonar takes acoustic measurements of the seabed and generates the sonar images while the UUV maintains a constant speed and heading. It is assumed that the UUV travels at a constant height and, thus, the vehicle has two degrees-of-freedom (DOF). It also has a minimum turning radius. It is assumed that the accurate measurements of UUV position and orientation are computed online via an onboard simultaneous localization and mapping (SLAM) algorithm described in [25]. Also, it is assumed that prior measurements of target positions in  $\mathcal{W}$  are known from prior sorties. Each target has a classification

label  $y \in \mathcal{Y}$ , where  $\mathcal{Y} = \{y_1, \dots, y_d\}$  is a finite set of possible target categories. The UUV obtains an observation  $o \in \Omega$  of the target class from a sonar image. The characteristics of sonar images are such that it is challenging to recognize target due to its low resolution and segmentation. Therefore, the UUV typically needs to acquire multiple observations from different aspect angles in order to estimate the features and classification of an underwater target. This thesis addresses the problem of classifying all the targets with satisfactory confidence using minimum time. The first step is to adaptively select the local viewpoint around each target and classify each target with sufficient confidence. The second step is to find an optimal global path that visits and classifies targets in minimum time.

## CHAPTER 3

### GENERAL POMDP FORMULATION

POMDP is an efficient approach for modeling sequential decision process in which decisions and transition probabilities obey the Markov property [26] [27]. Markov decision process (MDP) is defined by a collection of elements  $\{\mathcal{T}, \mathcal{S}, \mathcal{A}, P(s_{t+1}|s_t, a_t), R(s_t, a_t)\}$ , where each element is described as follows.  $\mathcal{T}$  is a set of decision epochs, where  $\mathcal{T} = \{1, 2, \dots, T\}, T \in \mathbb{N}$ .  $\mathcal{S}$  represents a finite set of possible system states and each system state is denoted as  $s \in \mathcal{S}$ .  $\mathcal{A}$  stands for a finite set of actions and its element is denoted as  $a \in \mathcal{A}$ . The transition probability is denoted by  $P(\cdot|s_t, a_t), \forall s_t \in \mathcal{S}, \forall a_t \in \mathcal{A}$ . For example,  $P(s_{t+1}|s_t, a_t)$  is the probability of the system states  $s_{t+1}$  at the next decision epoch  $t + 1$ , as a result of choosing action  $a_t$  in state  $s_t$  at current decision epoch  $t$ .  $R(s_t, a_t)$  is a reward that the agent receives as a result of choosing action  $a_t$  in state  $s_t$  at decision epoch  $t$ .

Since the system state  $s$  is often not fully observable, an extension of MDP, POMDP is introduced [28]. POMDP contains an observation set  $\Omega$  and observation model  $P(o_{t+1}|s_t, a_t, s_{t+1})$  in addition to the elements included in MDP. The observation model  $P(o_{t+1}|s_t, a_t, s_{t+1})$  stands for a probability distribution which determines the observation that the system receives at time epoch  $t + 1$ , given that the unobserved system occupies state  $s_t$  at decision epoch  $t$ , the agent chooses action  $a_t$ , and the unobserved system occupies state  $s_{t+1}$  at decision epoch  $t + 1$ . It is often assumed that the observation at next time epoch  $o_{t+1}$  is not depend on the current states  $s_t$ . Thus the observation model becomes  $P(o_{t+1}|a_t, s_{t+1})$ . Therefore, POMDP can be formulated by the collection of the elements  $\{\mathcal{T}, \mathcal{S}, \mathcal{A}, P(s_{t+1}|s_t, a_t), R(s_t, a_t), \Omega, P(o_{t+1}|a_t, s_{t+1})\}$ . This extension

enables POMDP to model the process whose true states are not observable.

To solve the partially observable problem, POMDP introduces the belief state  $b_t \in \mathcal{B}$ , which is the probability distribution over the possible states  $\mathcal{S}$  at time epoch  $t$ .  $b(s)$  is used to denote the probability of state  $s$  in the distribution  $b$ . The belief states  $b$  is updated by a new observation  $o_{t+1}$  as

$$\begin{aligned} b_{t+1}(s') &= P(s'|o_{t+1}, a_t, b_t) \\ &= \frac{P(o_{t+1}|a_t, s') \sum_{s \in \mathcal{S}} P(s'|a_t, s) b_t(s)}{P(o_{t+1}|a_t, b_t)}, \end{aligned} \quad (3.1)$$

where  $s' \in \mathcal{S}$  is an arbitrary state in state space  $\mathcal{S}$ . The action should be taken based on the reward function that can be estimated using current belief state, and therefore, the intermediate reward for taking action  $a_t$  in the belief state  $b_t$  is

$$\rho(a_t, b_t) = \sum_{s_t \in \mathcal{S}} b_t(s_t) R(a_t, s_t). \quad (3.2)$$

Using the belief state and the intermediate reward function, the POMDP can be solved similarly to the MDP formulated by  $\{\mathcal{T}, \mathcal{B}, \mathcal{A}, \rho(a_t, b_t), \tau(b_{t+1}, a_t, b_t)\}$ , where the transition function between belief states is formulated by

$$\begin{aligned} \tau(b_{t+1}, a_t, b_t) &= P(b_{t+1}|a_t, b_t) \\ &= \sum_{o_{t+1} \in \Omega} P(b_{t+1}|b_t, a_t, o_{t+1}) P(o_{t+1}|a_t, b_t). \end{aligned} \quad (3.3)$$

$P(b_{t+1}|b_t, a_t, o_{t+1}) = 1$  only if action  $a_t$  taken at belief state  $b_t$  obtaining observation  $o$  will lead to belief state  $b_{t+1}$ . The value iteration of the above MDP formulation aims to find a series of action which maximizes the sum of expected future value of the next  $m$  time steps,  $\mathbb{E}[\sum_{t=0}^{m-1} r_t \gamma^t]$ , where  $r_t$  is the reward received on time  $t$ .  $\gamma \in [0, 1]$  is a discount factor. The optimal value

function  $V_t^*$  at step  $t$  can be computed by

$$\begin{aligned}
V_t^*(b_t) &= \max_{a_t \in A} \mathbb{E} \left[ \sum_{t=0}^{m-1} r_t \gamma^t \right] \\
&= \max_{a_t \in A} \left\{ \rho(a_t, b_t) + \gamma \sum_{b_{t+1} \in \mathcal{B}} \tau(b_{t+1}, a_t, b_t) V_{t-1}^*(b_{t+1}) \right\} \\
&= \max_{a_t \in A} \left\{ \rho(a_t, b_t) + \gamma \sum_{o_{t+1} \in \Omega} P(o_{t+1} | a_t, b_t) V_{t-1}^*(b_{t+1}^{a_t, o_t}) \right\}, \tag{3.4}
\end{aligned}$$

where  $b_{t+1}^{a_t, o_t}$  is the belief states from action  $a_t$  and observation  $o_t$  starting from belief states  $b_t$ . The optimal action at any time step is the argument that maximizes the value function at that time step.



## CHAPTER 4

### IMAGE TEMPLATE ISSUE

In the paper [4] and [5], an image template  $M$  is generated for each possible target-aspect pair. In their work, there are four different object shape types, a cylinder, a truncated cone, a wedge-shaped object, and an irregularly shaped non-target object. In addition, the possible views are discretized into 8 aspects of 45 degrees each. Therefore, there are 32 image templates in total and the pixel size of them is  $119 * 460$ . As shown in the Figures 4.1 4.2, the image templates are clearly different from each other depend on their shape and view orientation. It is because the raw image data is acquired by synthetic aperture sonar (SAS) images and those images have high resolution.

In our sonar image data, the segmented images are divided into 8 different regions depend on the distance between the target and the vehicle and the relative angle  $\theta$  of the vehicle heading and target heading. Our targets have 2 types of shape, circle and square. Thus we should have 16 image templates in order to implement the same algorithm. However, our segmented images pixel size is too small to recognize the difference between different shape of targets as well as different regions. Figure 4.3 demonstrates 9 segmented images of a specific target (ID520) which has a rectangle shape, the pixel size of image is  $6 * 201$ . Figure 4.4 shows 4 segmented images of another target (ID517) which has a circle shape, the pixel size of image is  $11 * 151$ . Even though those two target have different shape, one is circle and the other is rectangle, all of those segmented images look the same here because of the low resolution. We can't really tell the difference between each regions as well. Therefore, we can't follow the algorithm in [4] to generate best matching image templates for each possible target-aspect

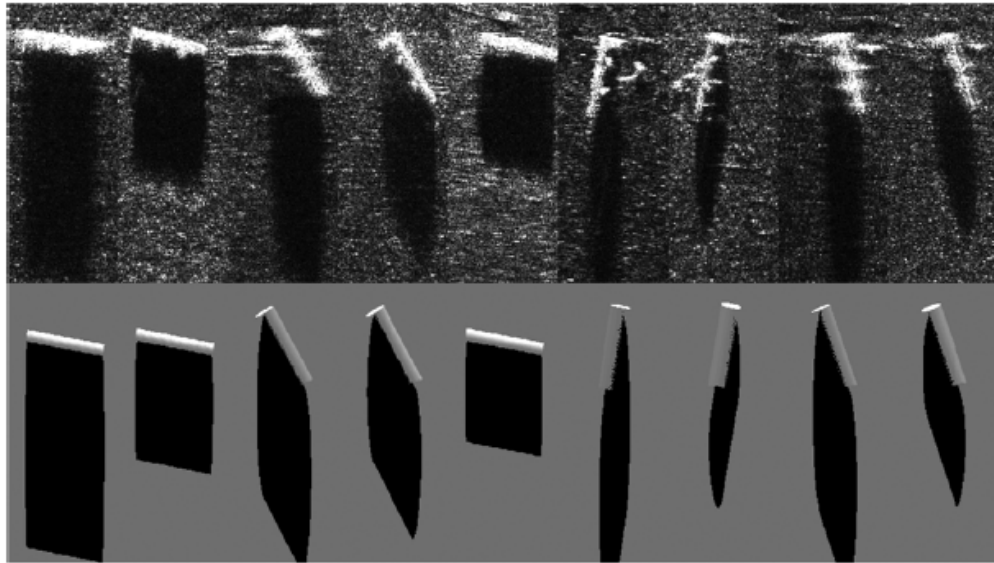


Figure 4.1: Cylinder target from several aspects and ranges with the corresponding matching templates (Myers, 2012)

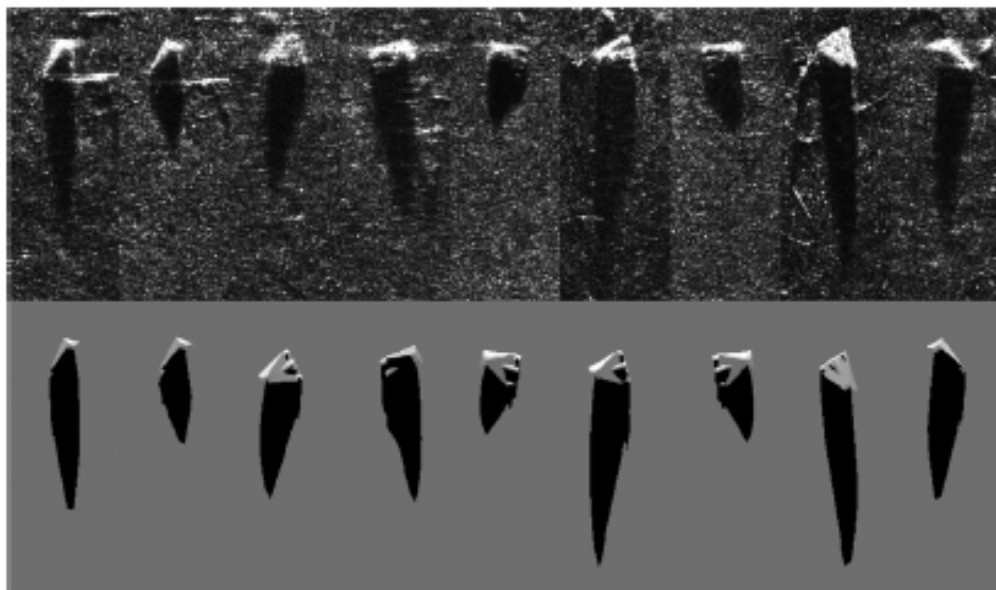


Figure 4.2: Wedge target from several aspects and ranges with the corresponding matching templates (Myers, 2012)

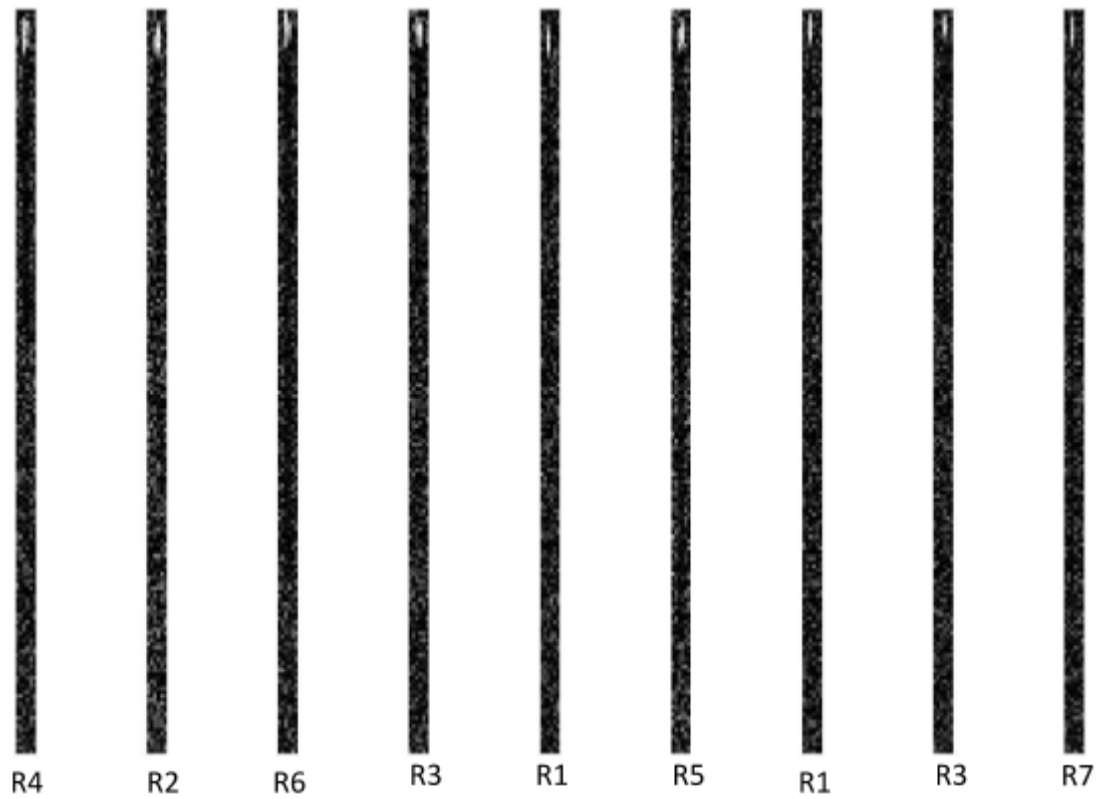


Figure 4.3: Segmented sonar images of rectangle shape target from several different regions

pair. Instead we need to figure out another way to get an observation model in the form of confidence level of target/non-target. The observation model is introduced in the next chapter.



Figure 4.4: Segmented sonar images of circle shape target from several different regions

### 5.1 Observation Model

In the general POMDP formulation, an observation model is required. In [5], the observation model of a classification POMDP is based on a template matching approach. An image template is generated for each possible target-aspect pair. However, a set of good templates are not always available for some applications. The performance of template matching can be affected by the resolution of the sensor images. To overcome this issue, we proposed an observation model that incorporates a deep Convolutional Neural Network (CNN) and Support Vector Machine (SVM). The proposed approach avoids the use of templates that need to be known to the robot *a-priori*. The proposed observation model can be easily adapted to other applications with different image sensors and environmental conditions.

The observation models is the probability of observing  $o \in \Omega$  given a current state and an action taken. These probabilities  $P(o_{t+1}|s_t, a_t, s_{t+1})$  are represented by a conditional probability table (CPT) learned from a training dataset by statistical analysis. Figure 5.1 shows the entire POMDP work flow.

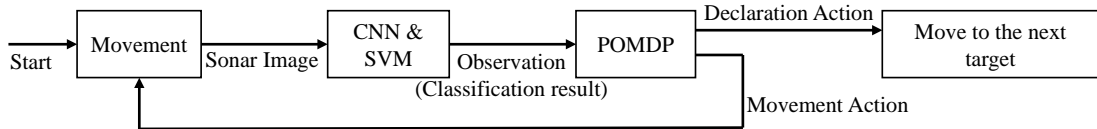


Figure 5.1: Work flow chart of PODMP-based underwater robot path planning

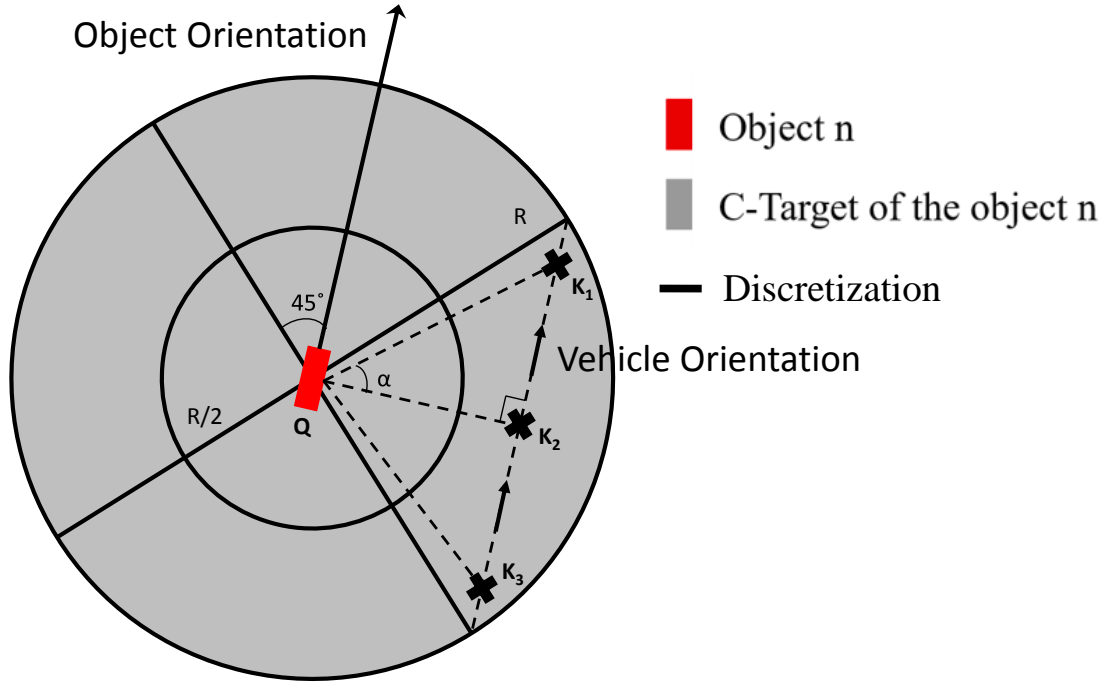


Figure 5.2: The discretization of the C-target based on the sensor range and aspect angle, which is defined as a region, and vehicle waypoints  $K_1, K_2, K_3$ .

## 5.2 C-target formation and Waypoints Selection

In compliance with the POMDP formulation, discretized robot configuration states are needed. This can be done by discretizing the C-target around each target of interest. C-target is defined as the set of configurations where the robot can observe the target. The C-target of every target is discretized into 8 regions of configurations  $\mathcal{C} = \{1, \dots, 8\}$  for all targets as shown in Figure 5.2. The 8 regions are generated by uniformly discretizing the aspect angle into four cases and the radius into two cases. The target orientation is the bisector of the region  $c = 1$ . During each POMDP process, the robot will inspect a single target. Thus, the robot's configuration is described using  $c$  only. In order for the onboard side-scan sonar to get good quality sonar image, the robot has to move in a straight line, which for the given C-target discretization, is either parallel or

perpendicular to the target orientation. Therefore, three waypoints have been created for the vehicle to explore each region as shown in the Figure 5.2.  $Q$  is the centroid of the target and the waypoint  $K_2$  is placed at the midpoint of the region in both distance and orientation.  $K_1K_3$  is perpendicular to  $QK_2$  and  $K_1, K_2, K_3$  three waypoints form a line segment as vehicle trajectory. The angles  $\angle K_1QK_2 = \angle K_2QK_3 = \alpha$ .

### 5.3 POMDP Supplement

For target classification and robot path planning, the system state is defined as a label-region pair  $s = (y, c)$  to describe the label  $y$  of the current inspecting target and robot's configuration  $c$ . There will be one absorbing end state  $\bar{s}$ . Once the system reaches the end state, the system remains in the end state with any actions. The set of possible state  $S$  is defined as follows,

$$S = \{(y, c) | y \in \mathcal{Y}, c \in \mathcal{C}\} \cup \bar{s}. \quad (5.1)$$

Since each region of configurations is large, it is assumed that the current configuration region  $c$  is known to the robot. Thus, the robot only needs to observe the target label from sonar images. The observation  $o$  only include the observed target label from sonar images via CNN and SVM.

The actions are divided into two types, 'movement' and 'declaration'. Movement type of action moves the vehicle into a certain region of the current target of interest. The set of movement type actions is  $\mathcal{A}_m$ . Each action in set  $\mathcal{A}_m$  moves the robot to one of the other regions of the current target. Declaration type of action declares the final classification label of the current target. Thus,

the set of actions is

$$\mathcal{A} = \mathcal{A}_m \cup \mathcal{A}_d. \quad (5.2)$$

In this specific problem, the transition model is defined as deterministic for both movement and declaration actions. For a movement action  $a \in \mathcal{A}_m$ , the vehicle obtains a view at the desired region with probability one. This model assumes that the vehicle can move to the exact commanded region around the target.

$$\begin{aligned} p(s_{t+1} = (y_n, c_{n,t+1}) | s = (y_n, c_{n,t}), a_t \in \mathcal{A}_m) \\ = \begin{cases} 1 & \text{if } c_{n,t+1} \equiv M(c_{n,t}, a_t) \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (5.3)$$

where  $c_{n,t}$  represents the configuration of the robot is at region  $c_n$  at time epoch  $t$ .  $M(c_{n,t}, a_t)$  is a function that outputs the next robot configuration region with current configuration region  $c_{n,t}$  and action  $a_t$ . A declaration action  $a \in \mathcal{A}_d$  declares the target classification and ends the POMDP process, which is modeled as entering into the absorbing end state  $\bar{s}$ .

The reward function is designed to treat all errors equally and not penalize the system for obtaining an additional view of the target as shown in Equation (5.4).

$$R(s, a) = \begin{cases} 1 & \text{Correct classification; } a \in \mathcal{A}_d, s \in S \\ -100 & \text{Incorrect classification; } a \in \mathcal{A}_d, s \in S \\ 0 & \text{Additional view; } a \in \mathcal{A}_m, s \in S \end{cases} \quad (5.4)$$

The discount factor  $\gamma$  is set to 0.75 to compute the value function of the model. The algorithm will computes sum of expected discounted reward of  $m = 5$  future steps.



## 5.4 Target ordering using TSP algorithm

The current POMDP algorithm will generate a policy for each target that consists of multiple movement actions and a declaration action in the end. The waypoints for the local path around each target is selected based on the policy of POMDP. Since each target is visited and classified once, the problem is formulated as a traveling salesman problem. Finding the optimal target visiting order is equivalent to solving the equivalent traveling salesman problem. This thesis uses the cross-entropy (CE) algorithm introduced in [23] and [24] to find an approximate solution to this TSP.

## CHAPTER 6

### NUMERICAL RESULTS

#### 6.1 CNN and SVM performance

To validate the proposed approach for robot path planning and demonstrate the performance of the proposed algorithm, the first step is to train a POMDP observation model from a pre-collected sonar image dataset obtained by an underwater robot equipped with a side-scan sonar. The observation set is denoted by  $\Omega = \{0, 1\}$  and  $y$  is the true classification label of each target. For each target in the training dataset, an observation label  $o$  from the combination of CNN and SVM is generated. The classification performance of the trained CNN and SVM is verified on a similar sonar image dataset. The results are shown in Table 6.1, where the elements of the confusion matrix are True Positive Rate (TPR), False Positive Rate (FPR), False Negative Rate (FNR), and True Negative Rate (TNR).

Table 6.1: Classification performance of the CNN and SVM confusion matrix

	$y = 1$	$y = 0$
$o = 1$	0.8837( <i>TPR</i> )	0( <i>FPR</i> )
$o = 0$	0.1163( <i>FNR</i> )	1( <i>TNR</i> )

The observation model  $P(o_{t+1}|s_t, a_t)$  is then learned from the dataset. Finally, the observation model is constructed and shown in Table 6.2.

Table 6.2: Conditionally probability table of observation model

<i>Region <math>r_m</math></i>		1	2	3	4	5	6	7	8
$y = 1$	$o = 0$	0.294	0.273	0.353	0.188	0.125	0.188	0.333	0.077
	$o = 1$	0.706	0.727	0.647	0.813	0.875	0.813	0.667	0.923
$y = 0$	$o = 0$	1	1	1	1	0.997	0.987	0.987	0.997
	$o = 1$	0	0	0	0	0.003	0.013	0.014	0.003

## 6.2 Observation Model Validation

The observation model is validated by comparing it with the probability distribution of randomly selected test data set. 30 objects (10 targets and 20 non-targets) are randomly selected from the total 520 objects for the simulation and those objects have 269 segmented sonar images. The same count method which gets the observation model is used here on the 269 test image data in order to get the conditional probability distribution and it is demonstrated in Table 6.3 and Table 6.4.

Table 6.3: The conditional probability table of test data when the ground truth  $y_i$  is 1 (target)

<i>Region <math>r_m</math></i>	1	2	3	4	5	6	7	8
$o = 0$	0.2222	0.25	0.4167	0.1	0.2	0.0769	0.4	0.1111
$o = 1$	0.7778	0.75	0.5833	0.9	0.8	0.9231	0.6	0.8889

Table 6.4: The conditional probability table of test data when the ground truth  $y_i$  is 0 (non-target)

<i>Region <math>r_m</math></i>	1	2	3	4	5	6	7	8
$o = 0$	1	1	1	1	1	1	0.95	1
$o = 1$	0	0	0	0	0	0	0.05	0

In order to prove the observation model works on the test data, we have to measure the similarity between the conditional probability in Table 6.2 and Table 6.3 for target, Table 6.4 for non-target. The chi-squared goodness of fit test is used to determine whether there is a significant difference between the expected frequencies and the observed frequencies in one or more categories [29] [30] [31]. It is the sum of differences between observed and expected outcome frequencies (that is, counts of observations), each squared and divided by the expectation:

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (6.1)$$

where  $n$  is the degree of freedom and equal to 1 since it is a binary case,  $E_i$  is the expectation value from the Table 6.2,  $O_i$  is the observation value from the Table 6.3 and Table 6.4.

Table 6.5: The chi square value when the ground truth  $y_i$  is 1 (target)

<i>Region <math>r_m</math></i>	1	2	3	4	5	6	7	8
$\chi^2$ for $o = 0$	0.0176	0.0019	0.0115	0.0408	0.045	0.0652	0.0133	0.0152
$\chi^2$ for $o = 1$	0.0073	0.0007	0.0063	0.0094	0.0064	0.015	0.0067	0.0013

Table 6.6: The chi square value when the ground truth  $y_i$  is 0 (non-target)

<i>Region <math>r_m</math></i>	1	2	3	4	5	6	7	8
$\chi^2$ for $o = 0$	0	0	0	0	0	0.0002	0.0013	0
$\chi^2$ for $o = 1$	0	0	0	0	0.003	0.0132	0.0985	0.0027

The results of the chi-square test are shown in the Table 6.5 and Table 6.6. If we set the confidence level to 95%, according to the chi-square possibilities

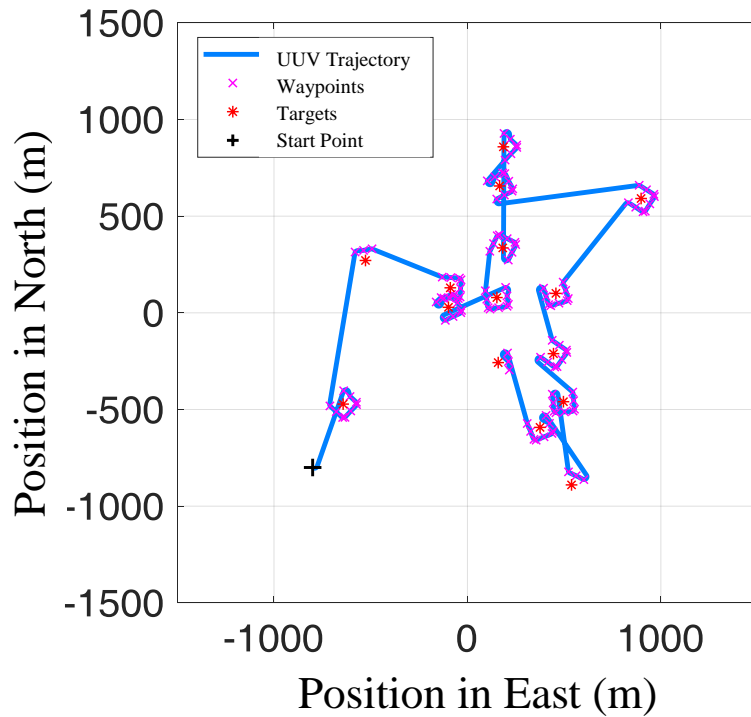
distribution table [32], the  $\chi^2$  value has to be less than 0.004. Therefore, 14 out of 16  $\chi^2$  values for the non-target objects in the Table 6.5 have passed the test, while in the Table 6.5 only 3 out of 16  $\chi^2$  values for the target objects have passed the test. The results demonstrates that the two conditional possibility distributions for the non-target objects are the same regardless of 5% noise level. So the observation model for the non-target object is relatively accurate.

However, for target objects, the chi-square test rejects most of the relationship between the test data possibility distribution and the observation model. The main reason is that the population of the data is too small. There are only 129 out of 3113 segmented sonar images of the target objects in the whole data set. For each of 8 regions, the number of the observation images is way less than the minimum sample size for the chi-square test. On the other hand, the sample size of the image of the non-target objects, which is 2984 out of 3113, is large enough for the chi-square test.

For small sample size such as the case of target objects, the Fisher's exact test is the more appropriate statistical significance test [33] [34]. The cross-validation techniques can also be implemented for both cases of target and non-target objects [35] [36]. It will be helpful to the validation and can also test whether the maximum likelihood observation model will perform well with new data [37] [38].

### 6.3 Simulation Results

The proposed path planning algorithm is implemented in a simulation environment including 15 targets. The ROI  $\mathcal{W}$  has a square shape with 3000-meter



side length. Ten of the targets are in class  $y = 0$  and the other five are in class  $y = 1$ . A sonar onboard robot is sent out to classify all the target using the proposed algorithm. The range of the sonar is 150 meters. For each target of interest, the initial belief state is set to be uniformly distributed over the state space  $\mathcal{S}$ , which indicates that the robot agent does not have prior information about the targets. The robot visited and classified targets within the ROI following the order generated by TSP solver. The robot will start the POMDP based classification algorithm from region  $c = 1$  of each target.

The algorithm generates a robot trajectory as shown in Figure 6.1 and the different zoom-in trajectories generated by POMDP policy for a single target are shown in Figure 6.2 Figure 6.3 Figure 6.4.

The performance is compared with other algorithms including multiple

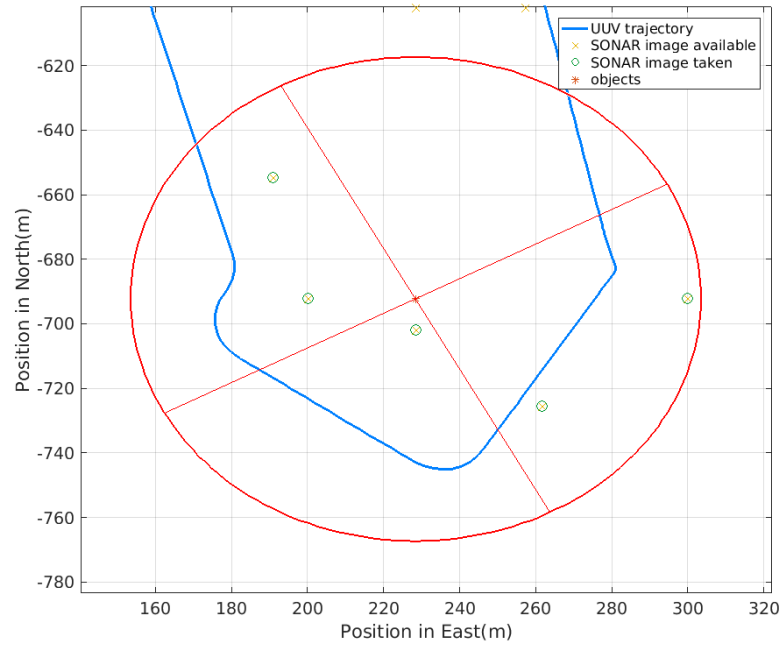


Figure 6.2: An example of vehicle paths for a POMDP policy in the data set. The regions visited are determined by the policy, and the waypoints are set to allow the vehicle to follow a straight line that will get a good quality sonar image.

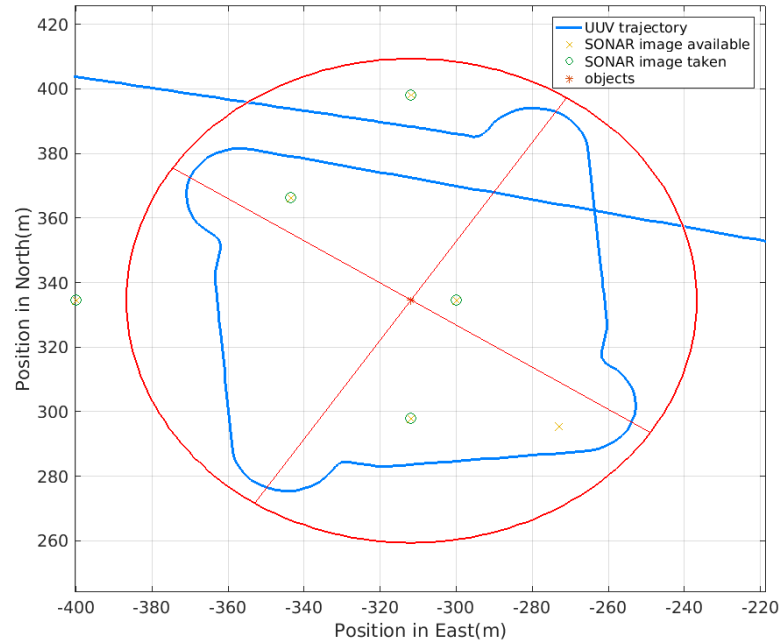


Figure 6.3: An example of vehicle paths for a POMDP policy in the data set.

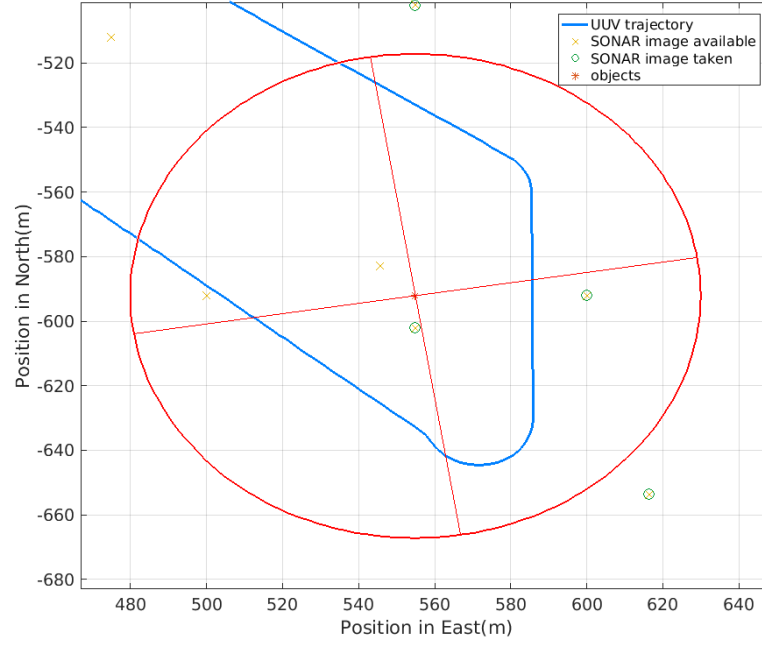


Figure 6.4: An example of vehicle paths for a POMDP policy in the data set.

aspect coverage (MAC) algorithm [39] and some of its extension. MAC algorithm solves the path planning problem as a coverage problem in several aspect angles by considering the estimated targets locations as a multivariate Gaussian distribution. For a large area, MAC algorithm can improve the efficiency by clustering nearby targets and visiting each cluster in the optimal order, as described in [40]. Figure 6.5 and Figure 6.6 shows the path generated by the MAC algorithm with or without clustering.

The classification and path planning performances are shown in Table 6.7. The results indicate that the proposed algorithm effectively acquire additional views to improve the classification accuracy and TSP provides an optimized path in the case of distance and travel time for the robot to visit each target. Although the proposed algorithm has similar good classification accuracy with other path planning algorithms for underwater robots, such as the MAC algorithms, the



Table 6.7: Performance comparison of the proposed algorithm and two MAC algorithms

Algorithm	Distance Traveled (m)	Time (h)	Average Number of Images per Target	TPR	FPR	Accuracy
POMDP and TSP	9754	1.32	3.33	0.8	0	0.93
MAC	61326	6.93	4.47	1	0	1
MAC with Clustering	42045	5.00	7.07	1	0	1

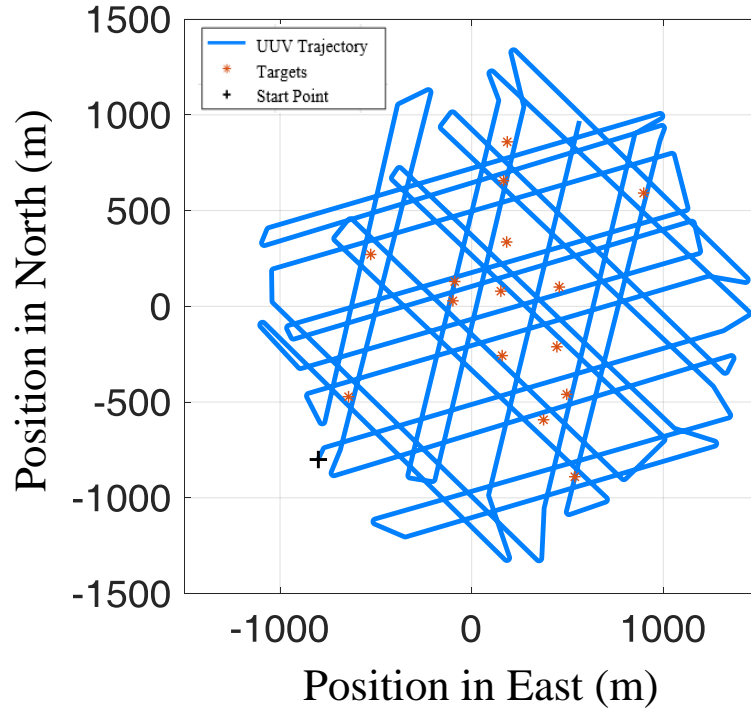


Figure 6.5: Robot trajectory generated by MAC algorithm

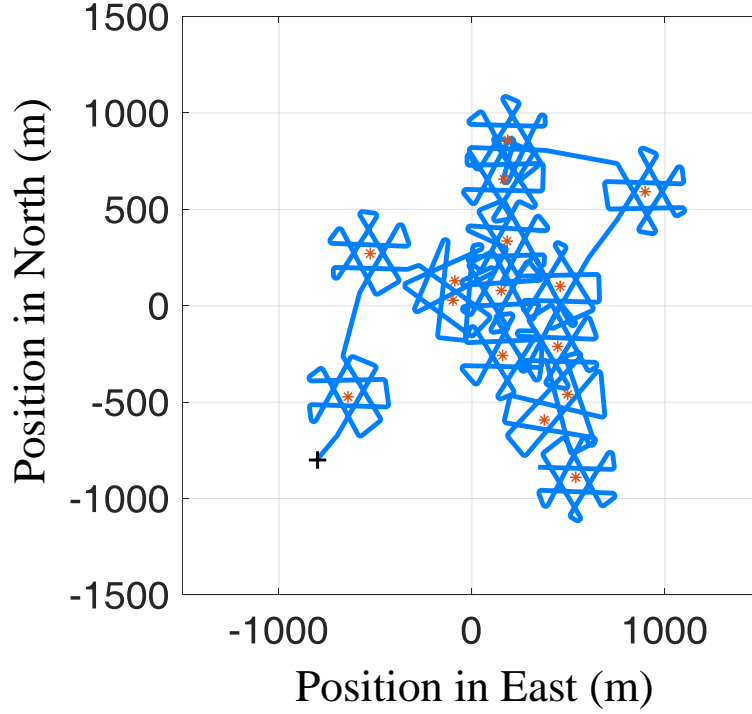


Figure 6.6: Robot trajectory generated by MAC algorithm with clustering

advantage of the POMDP and TSP algorithm is that this algorithm produces significantly shorter path while it still maintains satisfactory classification performance. Comparing to the MAC algorithm with or without clustering, the proposed algorithm find a path that is 76.8% and 84.1% shorter, respectively. Notice that the proposed algorithm has a true positive rate of  $TPR = 0.8$  that is lower than other MAC algorithms ( $TPR = 1$ ). One possible reason is that the number of targets in class  $y = 1$  is limited and much less than the number of targets in class  $y = 0$  in the current training dataset.

## CHAPTER 7

### CONCLUSION AND FUTURE WORK

This thesis introduces a multi-target multi-view classification algorithm that uses the image features extracted from sonar images by CNN and SVM as an observation model of a POMDP framework. The robot can move around the targets and select the region to visit adaptively in order to acquire additional views that most benefit the classification. The proposed algorithm decides which region should be visited next according to the reward function. A cross-entropy (CE) TSP algorithm is utilized to generate the optimal order of visiting all targets. Compared with state-of-the-art underwater robotic path planning algorithms, the numerical results demonstrate that the proposed algorithm can generate a much shorter path, while maintaining satisfactory performance in classification accuracy. The combination of the TSP and POMDP is an efficient solution of path planning.

One area of future work will be improving the classification rate of target class  $y = 1$ . Now the true positive rate of the proposed algorithm is not ideal comparing to other MAC algorithms. The other interesting direction of the future work is about the POMDP model for multi-target environment. Instead of using TSP to decide the target visiting order, the POMDP algorithm has the capability of making the decision to go to the configurations of different targets. The challenging part of multi-target POMDP framework is that the time complexity and the space complexity of the POMDP formulation is exponential to the number of states.

## BIBLIOGRAPHY

- [1] Xiaoguang Wang, Xuan Liu, Nathalie Japkowicz, and Stan Matwin. Automated approach to classification of mine-like objects using multiple-aspect sonar images. *Journal of Artificial Intelligence and Soft Computing Research*, 4(2):133–148, 2014.
- [2] David Williams and Johannes Groen. Multi-view target classification in synthetic aperture sonar imagery. In *Proceedings of the International Conference and Exhibition on Underwater Acoustic Measurements*, pages 699–704, 2009.
- [3] Marc Robinson, Mahmood R Azimi-Sadjadi, and Jaime Salazar. Multi-aspect target discrimination using hidden markov models and neural networks. *IEEE Transactions on Neural Networks*, 16(2):447–459, 2005.
- [4] Vincent Myers and John Fawcett. A template matching procedure for automatic target recognition in synthetic aperture sonar imagery. *IEEE Signal Processing Letters*, 17(7):683–686, 2010.
- [5] Vincent Myers and David P Williams. Adaptive multiview target classification in synthetic aperture sonar images using a partially observable markov decision process. *IEEE Journal of Oceanic Engineering*, 37(1):45–55, 2012.
- [6] Johan AK Suykens and Joos Vandewalle. Least squares support vector machine classifiers. *Neural processing letters*, 9(3):293–300, 1999.
- [7] Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, et al. A practical guide to support vector classification. 2003.
- [8] Pingping Zhu, Jason Isaacs, Bo Fu, and Silvia Ferrari. Deep learning feature extraction for target recognition and classification in underwater sonar images. In *Decision and Control (CDC), 2017 IEEE 56th Annual Conference on*, pages 2724–2731. IEEE, 2017.
- [9] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [10] Sven Behnke. *Hierarchical neural networks for image interpretation*, volume 2766. Springer, 2003.

- [11] Dan C Cireşan, Ueli Meier, Jonathan Masci, Luca M Gambardella, and Jürgen Schmidhuber. High-performance neural networks for visual object classification. *arXiv preprint arXiv:1102.0183*, 2011.
- [12] Dan C Ciresan, Ueli Meier, Jonathan Masci, Luca Maria Gambardella, and Jürgen Schmidhuber. Flexible, high performance convolutional neural networks for image classification. In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, volume 22, page 1237. Barcelona, Spain, 2011.
- [13] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [14] Patrice Y Simard, Dave Steinkraus, and John C Platt. Best practices for convolutional neural networks applied to visual document analysis. In *null*, page 958. IEEE, 2003.
- [15] Marti A. Hearst, Susan T Dumais, Edgar Osuna, John Platt, and Bernhard Scholkopf. Support vector machines. *IEEE Intelligent Systems and their applications*, 13(4):18–28, 1998.
- [16] Ingo Steinwart and Andreas Christmann. *Support vector machines*. Springer Science & Business Media, 2008.
- [17] Shi Chang, Jason Isaacs, Bo Fu, Jaejeong Shin, Pingping Zhu, and Silvia Ferrari. Confidence level estimation in multi-target classification problems. In *Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXIII*, volume 10628, page 1062818. International Society for Optics and Photonics, 2018.
- [18] Joseph B Kruskal. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical society*, 7(1):48–50, 1956.
- [19] Gregory Gutin and Abraham P Punnen. *The traveling salesman problem and its variations*, volume 12. Springer Science & Business Media, 2006.
- [20] Eugene L Lawler. The traveling salesman problem: a guided tour of combinatorial optimization. *Wiley-Interscience Series in Discrete Mathematics*, 1985.

- [21] Karla L Hoffman, Manfred Padberg, and Giovanni Rinaldi. Traveling salesman problem. In *Encyclopedia of operations research and management science*, pages 1573–1578. Springer, 2013.
- [22] Marco Dorigo and Luca Maria Gambardella. Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Transactions on evolutionary computation*, 1(1):53–66, 1997.
- [23] Reuven Rubinstein. The cross-entropy method for combinatorial and continuous optimization. *Methodology And Computing In Applied Probability*, 1(2):127–190, Sep 1999.
- [24] Pieter-Tjerk De Boer, Dirk P Kroese, Shie Mannor, and Reuven Y Rubinstein. A tutorial on the cross-entropy method. *Annals of operations research*, 134(1):19–67, 2005.
- [25] Hugh Durrant-Whyte and Tim Bailey. Simultaneous localization and mapping: part i. *IEEE robotics & automation magazine*, 13(2):99–110, 2006.
- [26] Qing Zhao, Lang Tong, Ananthram Swami, and Yunxia Chen. Decentralized cognitive mac for opportunistic spectrum access in ad hoc networks: A pomdp framework. *IEEE Journal on selected areas in communications*, 25(3), 2007.
- [27] Guy Shani, Joelle Pineau, and Robert Kaplow. A survey of point-based pomdp solvers. *Autonomous Agents and Multi-Agent Systems*, 27(1):1–51, 2013.
- [28] Guy Shani, Ronen I Brafman, and Solomon Eyal Shimony. Forward search value iteration for pomdps. In *IJCAI*, pages 2619–2624, 2007.
- [29] Ronald J Tallarida and Rodney B Murray. Chi-square test. In *Manual of Pharmacologic Calculations*, pages 140–142. Springer, 1987.
- [30] Albert Satorra and Peter M Bentler. A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66(4):507–514, 2001.
- [31] Henry Oliver Lancaster and Eugene Seneta. Chi-square distribution. *Encyclopedia of biostatistics*, 2, 2005.
- [32] Mary L McHugh. The chi-square test of independence. *Biochemia medica: Biochemia medica*, 23(2):143–149, 2013.

- [33] Michel Raymond and François Rousset. An exact test for population differentiation. *Evolution*, 49(6):1280–1283, 1995.
- [34] Peter Sprent. Fisher exact test. In *International encyclopedia of statistical science*, pages 524–525. Springer, 2011.
- [35] Ron Kohavi et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14, pages 1137–1145. Montreal, Canada, 1995.
- [36] Sylvain Arlot, Alain Celisse, et al. A survey of cross-validation procedures for model selection. *Statistics surveys*, 4:40–79, 2010.
- [37] Hirotugu Akaike. Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike*, pages 199–213. Springer, 1998.
- [38] Yoram Bresler and Albert Macovski. Exact maximum likelihood parameter estimation of superimposed exponential signals in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(5):1081–1089, 1986.
- [39] Matthew J Bays, Apoorva Shende, Daniel J Stilwell, and Signe A Redfield. Theory and experimental results for the multiple aspect coverage problem. *Ocean Engineering*, 54:51–60, 2012.
- [40] Abhishek Sriraman and Matthew J Bays. Efficient reacquire and identify path planning over large areas. In *Oceans-St. John's, 2014*, pages 1–7. IEEE, 2014.